Random Forest

m features

Dataset — n rows

RS+FS n' → M-1 Decision Tree 1

RS-R + FS-R → M-2 Decision Tree 2

RS-R + FS-R → M-3 Decision Tree 3

RS-R + FS-R → M-N Decision Tree 4
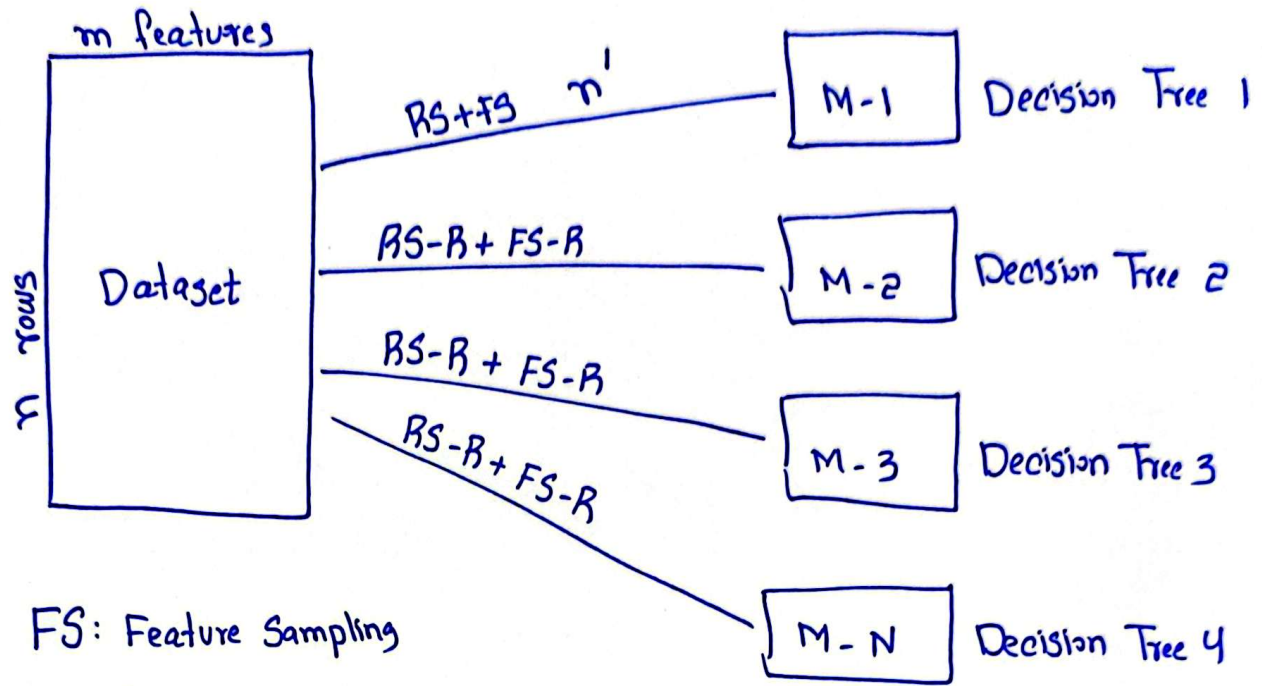
FS: Feature Sampling

RS: Row Sampling

FS-R: Feature Sampling with Replacement

RS-R: Row Sampling with Replacement

) new data

DT-1 → 1

DT-2 → 1

DT-3 → 0

DT-4 → 1

for classification

↳ majority voting classifier

$$[1,1,0,1] \Rightarrow 1$$

for regression

↳ average of all outputs

$$[1,1,0,1] \Rightarrow 0.75$$